

#2

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/511,794

Source: PCT

Date Processed by STIC: 10-27-04

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 10/27/2004

PATENT APPLICATION: US/10/511,794

TIME: 15:34:45

Input Set : A:\976-20 PCT US seq list.txt

Output Set: N:\CRF4\10272004\J511794.raw

3 <110> APPLICANT: Center for Genetic Engineering and Biotechnology  
 5 <120> TITLE OF INVENTION: ANTIBODY FRAGMENTS SPECIFIC FOR HUMAN CARCINOEMBRYONIC  
 6 ANTIGEN (CEA)  
 8 <130> FILE REFERENCE: 976-20 PCT/US  
 C--> 10 <140> **CURRENT APPLICATION NUMBER: US/10/511,794**  
 11 <141> CURRENT FILING DATE: 2004-10-19  
 13 <150> PRIOR APPLICATION NUMBER: PCT/CU2003/000005  
 14 <151> PRIOR FILING DATE: 2003-04-28  
 16 <150> PRIOR APPLICATION NUMBER: CU2002/0086  
 17 <151> PRIOR FILING DATE: 2002-04-09  
 19 <160> NUMBER OF SEQ ID NOS: 21  
 21 <170> SOFTWARE: PatentIn Ver. 2.1  
 23 <210> SEQ ID NO: 1  
 24 <211> LENGTH: 39  
 25 <212> TYPE: DNA  
 26 <213> ORGANISM: Artificial Sequence  
 28 <220> FEATURE:  
 29 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo  
 31 <400> SEQUENCE: 1  
 32 ggggatatcc accatgract tcgggytgag ctkggtttt 39  
 35 <210> SEQ ID NO: 2  
 36 <211> LENGTH: 29  
 37 <212> TYPE: DNA  
 38 <213> ORGANISM: Artificial Sequence  
 40 <220> FEATURE:  
 41 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo  
 43 <400> SEQUENCE: 2  
 44 ayctccacac acaggrccag tggatagac 29  
 47 <210> SEQ ID NO: 3  
 48 <211> LENGTH: 40  
 49 <212> TYPE: DNA  
 50 <213> ORGANISM: Artificial Sequence  
 52 <220> FEATURE:  
 53 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo  
 55 <400> SEQUENCE: 3  
 56 ggggatatcc accatggagw cacakwtca ggtctttrta 40  
 59 <210> SEQ ID NO: 4  
 60 <211> LENGTH: 21  
 61 <212> TYPE: DNA  
 62 <213> ORGANISM: Artificial Sequence  
 64 <220> FEATURE:  
 65 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo  
 67 <400> SEQUENCE: 4

## RAW SEQUENCE LISTING

DATE: 10/27/2004

PATENT APPLICATION: US/10/511,794

TIME: 15:34:45

Input Set : A:\976-20 PCT US seq list.txt

Output Set: N:\CRF4\10272004\J511794.raw

```

68 actggatggt gggaagatgg a 21
71 <210> SEQ ID NO: 5
72 <211> LENGTH: 14
73 <212> TYPE: PRT
74 <213> ORGANISM: Artificial Sequence
76 <220> FEATURE:
77 <223> OTHER INFORMATION: Description of Artificial Sequence: linker I
79 <400> SEQUENCE: 5
80 Glu Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Val Asp
81 1 5 10
84 <210> SEQ ID NO: 6
85 <211> LENGTH: 5
86 <212> TYPE: PRT
87 <213> ORGANISM: Artificial Sequence
89 <220> FEATURE:
90 <223> OTHER INFORMATION: Description of Artificial Sequence: linker II
92 <400> SEQUENCE: 6
93 Gly Gly Gly Gly Ser
94 1 5
97 <210> SEQ ID NO: 7
98 <211> LENGTH: 39
99 <212> TYPE: DNA
100 <213> ORGANISM: Artificial Sequence
102 <220> FEATURE:
103 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo
105 <400> SEQUENCE: 7
106 tctcacagtg cacaggaagt gaagctggtg gagtctggg 39
109 <210> SEQ ID NO: 8
110 <211> LENGTH: 63
111 <212> TYPE: DNA
112 <213> ORGANISM: Artificial Sequence
114 <220> FEATURE:
115 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo
117 <400> SEQUENCE: 8
118 gtcgactttg gattcggagc ctgatactga ggatttacct tctgaggaga ctgtgagagt 60
119 ggt 63
122 <210> SEQ ID NO: 9
123 <211> LENGTH: 62
124 <212> TYPE: DNA
125 <213> ORGANISM: Artificial Sequence
127 <220> FEATURE:
128 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo
130 <400> SEQUENCE: 9
131 gagggtaaatt cctcaggatc aggtctccgaa tccaaagtcg acgacattgt gatgacccag 60
132 tc 62
135 <210> SEQ ID NO: 10
136 <211> LENGTH: 37
137 <212> TYPE: DNA
138 <213> ORGANISM: Artificial Sequence

```

## RAW SEQUENCE LISTING

DATE: 10/27/2004

PATENT APPLICATION: US/10/511,794

TIME: 15:34:45

Input Set : A:\976-20 PCT US seq list.txt

Output Set: N:\CRF4\10272004\J511794.raw

```

140 <220> FEATURE:
141 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo
143 <400> SEQUENCE: 10
144 aaggaaaaaa gcggccgctt tcagctccag cttgggtt          37
147 <210> SEQ ID NO: 11
148 <211> LENGTH: 36
149 <212> TYPE: DNA
150 <213> ORGANISM: Artificial Sequence
152 <220> FEATURE:
153 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo
155 <400> SEQUENCE: 11
156 agagccgccg ccacctgagg agactgtgag agtggt          36
159 <210> SEQ ID NO: 12
160 <211> LENGTH: 36
161 <212> TYPE: DNA
162 <213> ORGANISM: Artificial Sequence
164 <220> FEATURE:
165 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo
167 <400> SEQUENCE: 12
168 ggtggcggcg gctctgacat tgtgatgacc cagtct          36
171 <210> SEQ ID NO: 13
172 <211> LENGTH: 108
173 <212> TYPE: DNA
174 <213> ORGANISM: Artificial Sequence
176 <220> FEATURE:
177 <223> OTHER INFORMATION: Description of Artificial Sequence: vector
179 <400> SEQUENCE: 13
180 cttttctatt ctacagtgc acaggaaatc aaagcggccg cagggtccga acaaaaactc 60
181 atctcagaag aggatctgaa ttcccatcat catcaccatc actaataa          108
184 <210> SEQ ID NO: 14
185 <211> LENGTH: 21
186 <212> TYPE: DNA
187 <213> ORGANISM: Artificial Sequence
189 <220> FEATURE:
190 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo
192 <400> SEQUENCE: 14
193 gttgttcctt tctattctca c          21
196 <210> SEQ ID NO: 15
197 <211> LENGTH: 24
198 <212> TYPE: DNA
199 <213> ORGANISM: Artificial Sequence
201 <220> FEATURE:
202 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo
204 <400> SEQUENCE: 15
205 ctcttctgag atgagttttt gttc          24
208 <210> SEQ ID NO: 16
209 <211> LENGTH: 241
210 <212> TYPE: PRT
211 <213> ORGANISM: Artificial Sequence

```

## RAW SEQUENCE LISTING

DATE: 10/27/2004

PATENT APPLICATION: US/10/511,794

TIME: 15:34:45

Input Set : A:\976-20 PCT US seq list.txt

Output Set: N:\CRF4\10272004\J511794.raw

213 &lt;220&gt; FEATURE:

214 &lt;223&gt; OTHER INFORMATION: Description of Artificial Sequence: scFv

216 &lt;400&gt; SEQUENCE: 16

217 Glu Val Lys Leu Val Glu Ser Gly Gly Asp Leu Val Lys Pro Gly Gly

218 1 5 10 15

220 Ser Leu Lys Phe Ser Cys Ala Ala Ser Gly Phe Pro Phe Asn Arg Tyr

221 20 25 30

223 Ala Met Ser Trp Val Arg Gln Thr Pro Glu Lys Arg Leu Glu Trp Val

224 35 40 45

226 Ala Phe Ile Ser Ser Asp Gly Ile Ala Tyr Tyr Ala Asp Ser Val Lys

227 50 55 60

229 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Arg Asn Ile Leu Tyr Leu

230 65 70 75 80

232 Gln Met Ser Ser Leu Arg Ser Glu Asp Thr Ala Met Tyr Tyr Cys Ala

233 85 90 95

235 Arg Val Tyr Tyr Tyr Gly Ser Ser Tyr Phe Asp Tyr Trp Gly Gln Gly

236 100 105 110

238 Thr Thr Leu Thr Val Ser Ser Glu Gly Lys Ser Ser Gly Ser Gly Ser

239 115 120 125

241 Glu Ser Lys Val Asp Asp Ile Val Met Thr Gln Ser Pro Lys Phe Met

242 130 135 140

244 Ser Thr Ser Val Gly Asp Arg Val Ser Val Thr Cys Lys Ala Ser Gln

245 145 150 155 160

247 Asn Ala Gly Thr Asn Val Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ser

248 165 170 175

250 Pro Lys Ala Leu Ile Tyr Ser Ala Ser Ser Arg Asn Ser Gly Val Pro

251 180 185 190

253 Asp Arg Ile Thr Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile

254 195 200 205

256 Ser Asn Val Gln Ser Glu Asp Leu Ala Glu Tyr Phe Cys Gln Gln Tyr

257 210 215 220

259 Asn Ser Tyr Pro Leu Val Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu

260 225 230 235 240

262 Lys

266 &lt;210&gt; SEQ ID NO: 17

267 &lt;211&gt; LENGTH: 232

268 &lt;212&gt; TYPE: PRT

269 &lt;213&gt; ORGANISM: Artificial Sequence

271 &lt;220&gt; FEATURE:

272 &lt;223&gt; OTHER INFORMATION: Description of Artificial Sequence: diabody

274 &lt;400&gt; SEQUENCE: 17

275 Glu Val Lys Leu Val Glu Ser Gly Gly Asp Leu Val Lys Pro Gly Gly

276 1 5 10 15

278 Ser Leu Lys Phe Ser Cys Ala Ala Ser Gly Phe Pro Phe Asn Arg Tyr

279 20 25 30

281 Ala Met Ser Trp Val Arg Gln Thr Pro Glu Lys Arg Leu Glu Trp Val

282 35 40 45

284 Ala Phe Ile Ser Ser Asp Gly Ile Ala Tyr Tyr Ala Asp Ser Val Lys

285 50 55 60

## RAW SEQUENCE LISTING

DATE: 10/27/2004

PATENT APPLICATION: US/10/511,794

TIME: 15:34:45

Input Set : A:\976-20 PCT US seq list.txt

Output Set: N:\CRF4\10272004\J511794.raw

```

287 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Arg Asn Ile Leu Tyr Leu
288 65          70          75          80
290 Gln Met Ser Ser Leu Arg Ser Glu Asp Thr Ala Met Tyr Tyr Cys Ala
291          85          90          95
293 Arg Val Tyr Tyr Tyr Gly Ser Ser Tyr Phe Asp Tyr Trp Gly Gln Gly
294          100         105         110
296 Thr Thr Leu Thr Val Ser Ser Gly Gly Gly Gly Ser Asp Ile Val Met
297          115         120         125
299 Thr Gln Ser Pro Lys Phe Met Ser Thr Ser Val Gly Asp Arg Val Ser
300          130         135         140
302 Val Thr Cys Lys Ala Ser Gln Asn Ala Gly Thr Asn Val Ala Trp Tyr
303 145          150         155         160
305 Gln Gln Lys Pro Gly Gln Ser Pro Lys Ala Leu Ile Tyr Ser Ala Ser
306          165         170         175
308 Ser Arg Asn Ser Gly Val Pro Asp Arg Ile Thr Gly Ser Gly Ser Gly
309          180         185         190
311 Thr Asp Phe Thr Leu Thr Ile Ser Asn Val Gln Ser Glu Asp Leu Ala
312          195         200         205
314 Glu Tyr Phe Cys Gln Gln Tyr Asn Ser Tyr Pro Leu Val Thr Phe Gly
315          210         215         220
317 Ala Gly Thr Lys Leu Glu Leu Lys
318 225          230
321 <210> SEQ ID NO: 18
322 <211> LENGTH: 35
323 <212> TYPE: DNA
324 <213> ORGANISM: Artificial Sequence
326 <220> FEATURE:
327 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo
329 <400> SEQUENCE: 18
330 catgccatgg ggaatccgaa gtgaagctgg tggag 35
333 <210> SEQ ID NO: 19
334 <211> LENGTH: 36
335 <212> TYPE: DNA
336 <213> ORGANISM: Artificial Sequence
338 <220> FEATURE:
339 <223> OTHER INFORMATION: Description of Artificial Sequence: oligo
341 <400> SEQUENCE: 19
342 catgccatgg atcccgggggt gatggtgatg gtgatg 36
345 <210> SEQ ID NO: 20
346 <211> LENGTH: 21
347 <212> TYPE: DNA
348 <213> ORGANISM: Artificial Sequence
350 <220> FEATURE:
351 <223> OTHER INFORMATION: Description of Artificial Sequence: diabody MS
353 <400> SEQUENCE: 20
354 gactggttcc aattgacaag c 21
357 <210> SEQ ID NO: 21
358 <211> LENGTH: 255
359 <212> TYPE: PRT

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/511,794

DATE: 10/27/2004

TIME: 15:34:46

Input Set : A:\976-20 PCT US seq list.txt

Output Set: N:\CRF4\10272004\J511794.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number